

Amendments to the Specification

Please replace the paragraph beginning at page 6, line 22, with the following rewritten paragraph:

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990, *J. Mol. Biol.* 215:403-10) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. ~~It can be accessed at <http://www.ncbi.nlm.nih.gov/BLAST/>. A description of how to determine sequence identity using this program is available at http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.~~

Please replace the paragraph beginning at page 6, line 28, with the following rewritten paragraph:

Homologs of the disclosed 15 kDa selenoprotein are typically characterized by possession of at least 70% sequence identity counted over the full length alignment with the amino acid sequence of a selected transcription factor using the NCBI Blast 2.0, Basic BLAST search, gapped blastp program set to default parameters (BLOSUM62 matrix; Gap existence cost=11; Per residue gap cost=1; lambda ratio=0.85). Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 75%, at least 80%, at least 90% or at least 95% sequence identity. When less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 75% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or 95% depending on their similarity to the reference sequence. ~~Methods for determining sequence identity over such short windows are described at http://www.ncbi.nlm.nih.gov/BLAST/blast_FAQs.html.~~ One of skill in the art will appreciate that these sequence identity ranges are provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided. The present invention provides not only the peptide homologs as described above, but also nucleic acid molecules that encode such homologs.

Please replace the paragraph beginning at page 12, line 25, with the following rewritten paragraph:

The human 15 kDa selenoprotein was detected in and purified from the human Jurkat T-cell line, JPX9 (Nagata et al., 1989, *J. Virol.* 63:3220-6) by growing the cells in the presence of ^{75}Se followed by analysis of extracts of the ^{75}Se -labeled cells by SDS PAGE and ~~PhosphorImager~~ PHOSPHORIMAGER detection of radioactivity on the gels. One of the major ^{75}Se -labeled proteins that migrated as a 15 kDa band on SDS PAGE was purified initially on DEAE-Sepharose and phenyl-Sepharose columns, and then further on a reverse-phase column. The procedures used were as follows. JPX9, was grown and labeled with [^{75}Se]selenious acid (2 $\mu\text{Ci/ml}$) as described in Gladyshev et al. (*Proc. Natl. Acad. Sci. USA* 93:6146-51, 1996). ^{75}Se -labeled JPX9 cells were mixed with unlabeled cells, suspended in 2 volumes of 30 mM Tris-HCl, pH 7.5, 1 mM EDTA, 2 mM DTT, 1 mM MgCl_2 , 1 mM phenylmethylsulfonyl fluoride and disrupted by sonication. Disrupted cells were centrifuged, the supernatant applied to a DEAE-Sepharose column, which had been equilibrated with 30 mM Tris-HCl, pH 7.5, 2 mM DTT and 1 mM EDTA (buffer A), the column washed with 2 volumes of buffer A and proteins eluted by application of a linear gradient from 0 to 500 mM NaCl in buffer A. Fractions containing ^{75}Se were analyzed on SDS gels. Fractions containing the human 15 kDa selenoprotein that eluted from the DEAE column with 350 mM NaCl were combined, concentrated, adjusted to a concentration of 0.5 M ammonium sulfate in buffer A, applied to a phenyl-Sepharose column equilibrated in 1 M ammonium sulfate in buffer A, the column washed by application of a linear gradient from 0.5 to 0 M ammonium sulfate in buffer A, and radioactive fractions corresponding to the 15 kDa selenoprotein eluted by application of a linear gradient from buffer A to water. Radioactive fractions were combined, concentrated, and loaded on a C_{18} reverse-phase HPLC column that had been equilibrated in 0.05% trifluoroacetic acid (TFA), a gradient of 0 to 60% acetonitrile in 0.05% TFA applied and ^{75}Se -containing fractions corresponding to the 15 kDa selenoprotein eluted at 48% acetonitrile.

Please replace the paragraph beginning at page 14, line 8, with the following rewritten paragraph:

The sequences of three different tryptic peptides and one overlapping peptide from the 15 kDa selenoprotein were determined. Computer searches of the partial cDNAs in the expressed sequence tags database (dbEST) using TBLASTN program revealed nucleotide sequences that corresponded to all three peptides in the same ORF. These cDNA sequences were used to assemble an open reading frame, depicted in FIG. 1. The two cDNA clones containing the longest 5' sequences were obtained from I.M.A.G.E. Consortium at the Lawrence Livermore National Laboratory (California) (<http://www.bio.llnl.gov/obrp/image/image.html>) and sequenced. These clones revealed a continuous nucleotide sequence of 1268 nucleotides, containing a single open reading frame of 162 amino acid residues and a 3'-end polyA tail. A single ATG codon occurs in a nucleotide context, GCGATGG, that is similar to the Kozak consensus sequence for initiation of translation (Kozak, 1997, *EMBO J.* 16:2482-92). This initiation ATG codon is followed by a 489 nucleotide open reading frame with an in-frame TAA termination codon. The obtained ORF includes an in-frame TGA codon, suggesting the presence of a selenocysteine residue, Sec93. Three tryptic peptides for which sequences have been determined correspond to deduced sequences located downstream of the TGA codon, indicating readthrough of the TGA codon, rather than termination of translation. Although selenocysteine was not directly identified as a component of the 15 kDa selenoprotein, the labeling of the protein with ⁷⁵Se, readthrough of the TGA codon and the location of selenocysteine insertion sequence (SECIS) element in the untranslated area (below) suggest the presence of selenocysteine in the protein. The predicted ORF encoded a protein of 17,790.6 Da. The mass of the purified 15 kDa selenoprotein was 14,870 Da, and this discrepancy suggested post-translational processing of the protein. Processing of the 15 kDa selenoprotein appears to occur at the N-terminal portion of the protein. Since antiserum raised to a synthetic peptide that was identical in sequence to the eighteen C-terminal residues of the 15 kDa selenoprotein, it recognized the 15 kDa selenoprotein at different stages of purification. In addition, one of the sequenced tryptic peptides obtained from digests of the 15 kDa selenoprotein corresponded to residues 146-158, located near the C-terminus according to the predicted gene sequence.

Please replace the paragraph beginning at page 15, line 23, with the following rewritten paragraph:

The amino acid sequence of the mouse protein was deduced from the assembly of 39 independent partial cDNA sequences in dbEST. In addition, experimental confirmation of the 5' region encoding the mouse N-terminal sequence was made from partial cDNAs obtained from the IMAGE consortium. The *C. elegans* sequence was assembled from two partial cDNA clones (GenBank dbEST accession numbers C10051 and C08344) which are identical for an 81 bp region of overlap and encode the apparently complete reading frame shown. The partial amino acid sequence of the homolog from the filarial nematode, *B. malayi*, was translated from a single partial cDNA (GenBank dbEST accession number AA257328). Two rice partial cDNAs (GenBank dbEST accession numbers D47693, D47819) covered the translated region shown (in addition, shorter segments of similarity to the human sequence were noted in translations further downstream, but these were in error-prone regions of mismatch between the two ESTs and are not shown). All pairwise alignments were strongly significant, as shown by TBLASTX-2 (Washington University gapped blast, February 1997 release ~~obtained from~~ <http://blast.wustl.edu/blast/executables>). Typical EST pairs gave amino acid gapped E (expect) values (BLOSUM 62 matrix), using the sum statistics of Altschul and Gish (*Methods Enzymol.* 266:460-80, 1996) as follows (with the highest HSP score appended in parentheses): human/mouse: 2×10^{-35} [717]; human/*C. elegans*: 2×10^{-20} [252]; human/*B. malayi*: 8×10^{-12} [228]; *C. elegans*/*B. malayi*: 8×10^{-21} [257]; human/rice (including multiple short matches for scoring purposes): 1×10^{-2} [82].